Microarray-based characterization of microbial community functional structure and heterogeneity in Sediments from the Gulf of Mexico

Liyou Wu[1,2,3], Laurie L.E. Kellogg[1], C.W. Schadt[2], Allan H. Devol[4], James M. Tiedje[5], and Jizhong Zhou[1,2,3]

[1]University of Oklahoma, Norman, OK, [2]Oak Ridge National Laboratory, Oak Ridge, TN, [3]Virtual Institute for Microbial Stress and Survival, Berkeley, CA, [4]University of Washington, Seattle, WA, [5]Michigan State University, East Lasing, MI

Q-381 | VIMSS Virtual Institute for Microbial Stress and Survival | University of Washington | IEG Institute for Environmental Genomics | Oak Ridge National Laboratory | Michigan State University | http://ieg.ou.edu/

ABSTRACT

Marine sediments of coastal margins are the dominant sites of carbon sequestration and nitrogen cycling. To determine the metabolic potential and structure of marine sediment microbial communities, 12 marine sediment samples from different sediment depths of two stations (GMT at 200m and GMS at 800m) in the Gulf of Mexico were analyzed using the functional gene arrays (FGAs) containing ~2000 probes from the genes involved in carbon fixation, organic carbon degradation, contaminant degradation, metal resistance, and nitrogen, sulfur, and phosphorous cycling. The geochemistry was highly variable for the sediments based on both site and depth. High microbial functional gene diversity was observed in these marine sediment samples. A total of 1343 **(68.5%)** genes of the probes belonging to various functional gene categories showed significant hybridization with at least one of the 12 samples. Overall functional gene diversity of samples from shallow depths was, in general, higher than those from deep depths at both stations; and the functional gene diversity of the samples at GMS was generally higher than that at GMT. High microbial heterogeneity existed in these marine sediments. In general, the microbial community structure was more similar when the samples were spatially closer. The number of unique genes at GMT increased with depth, from 1.7% at 0.75 cm to 18.9% at 25 cm depth. The same trend occurred at GMS, from 1.2% at 0.25 cm to 15.2% at 16 cm. In addition, the marine sediments appeared to possess high metabolic potential related to various biogeochemical processes such as carbon degradation, nitrification, denitrification, nitrogen fixation, sulfur reduction, phosphorus utilization and contaminant degradation. Finally, the Mantel test revealed significant positive correlations between various specific functional genes and functional processes although no overall correlation was observed among the geochemistry variables we measured with the microbial functional group communities. These results suggest that the functional gene arrays are a powerful tool for analyzing microbial community functional structure, these marine sediment microbial communities are diverse and spatially heterogeneous, and that the relationships between community structure and biogeochemical measurements appear to be complicated.

SAMPLES

Marine sediment samples were collected from two stations of the Gulf of Mexico: Station 3 (GMT), the 200 m water column depth (latitude, 26° 41.87' North; longitude, 96° 24.913' West) and Station 6 (GMS), the 800 m water column depth (latitude, 26° 40.37' North; longitude, 96° 06.84' West). Sediment cores, with overlying water, were collected with a Soutar box core and subcores at different sediment depths were taken using 7.5 and 10 cm cast-acrylic tubes.

Table 1. Geochemistry measured in each depth increment by Station

Stations	Sample code	Sediment Depth (cm)[a]	PO_4[b]	NO_3+NO_2[b]	NH_4[b]	SO_4[b]	Mn (II)[b]	Porosity	$Si(OH)_4$[b]	O_2[b]
#3 (200 m)	GMT-1	0.75	1.62	20.07	5.20	28.5	nd[c]	0.899	111.3	0
	GMT-3	2.50	6.05	-0.04	8.86	28.4	21.0	0.869	93.1	0
	GMT-5	4.50	7.20	0.40	27.21	28.4	22.6	0.849	133.6	0
	GMT-7	9.00	6.16	0.04	39.22	28.5	20.5	0.827	117.2	0
	GMT-9	16.00	7.25	0.04	38.35	28.5	19.4	0.785	154.5	0
	GMT-11	25.50	7.03	0.11	110.42	28.6	20.3	0.798	111.3	0
#6 (800 m)	GMS-1	0.25	4.27	34.90	11.46	28.8	nd[c]	0.907	85.4	112.8
	GMS-3	1.25	6.34	25.20	3.77	28.8	2.6	0.871	127.7	55.1
	GMS-5	2.50	9.65	2.08	2.28	28.8	64.7	0.863	149.5	15.4
	GMS-7	4.50	5.98	nd	17.04	28.7	137.1	0.847	125.9	0
	GMS-9	9.00	12.08	0.00	44.76	28.6	214.8	0.849	185.9	0
	GMS-11	16.00	15.80	0.00	116.03	28.8	nd[c]	0.843	250.0	0

[a] Data is shown only for sediment depths for which microarray analyses were performed
[b] Concentrations are in μM
[c] Not determined

METHODS

- **DNA extraction.** The bulk community DNA was directly extracted from 2 g of each sediment sample.
- **50mer Oligonucleotide functional gene array construction:** The functional gene array (FGA-I) was constructed using a diverse set of functional genes which are shown in table 2.

Table 2. The composition of the probes used in the FGA I

Functional groups	No. of gene categories[a]	No. Probes[b]
Carbon degradation	8	422
Carbon fixation	4	131
Contaminant degradation	96	731
Phosphorous metabolism	4	81
Nitrogen cycle	13	319
Sulfur cycle	2	204
Metal resistance	1	81
Misc	1	2

a Orthologs of the same gene are only counted once.
b Each probe is complementary to one ortholog of a given gene.

- DNA amplification, labeling, and hybridization. 100 ng of DNA of each sample was amplified using phi29 DNA polymerase, labeled with cy5 and hybridized to the FGA I.
- Microarray scanning and data processing. Hybridized microarray slides were scanned using a ScanArray® 5000 and the image displays were analyzed by quantifying the pixel density (intensity) of each spot using ImaGene™ version 5.0. Empty and poor spots were removed before the signal intensities were normalized by the mean signal across the slide; then outliers (at p<0.01) and minorities (only 1of the three replicates was present) were also removed.
- Data Analysis. Functional gene diversity was calculated using Simpson's reciprocal index (1/D) and Shannon-Weaver index (H). Cluster analysis was performed using the pairwise average-linkage hierarchical clustering algorithm in the CLUSTER software. Several multivariate statistical methods, Mantel test, DCA and CCA analysis, were employed to analyze the microarray data.

RESULTS: Gene diversity and gene overlaps

Table 3. Gene **Overlap** (Un-shaded Percentages) between Depth, Gene **Uniqueness** (Shaded Percentages), and **Diversity** Indices for Station 3

Depth (cm)	GMT-1	GMT-3	GMT-5	GMT-7	GMT-9	GMT-11
GMT-1	1.70%	66.40%	70.30%	63%	50%	44.70%
GMT-3		4.60%	65.90%	65.30%	58.90%	54.40%
GMT-5			1.90%	67.70%	57.50%	50.90%
GMT-7				1.80%	62.10%	59.80%
GMT-9					12.10%	63.80%
GMT-11						18.90%
Total genes detected	291	373	324	394	522	593
Simpson's (1/D)	118.4	133.1	134.6	136.4	168.2	201.6
Shannon Weaver's H'	5.185	5.266	5.175	5.203	5.311	5.282
Shannon Weaver's evenness	0.797	0.809	0.795	0.8	0.816	0.812

Table 4. Gene Overlap (Un-shaded Percentages) between Depth, Gene Uniqueness (Shaded Percentages), and Diversity Indices for Station 3

Depth	GMS-1	GMS-3	GMS-5	GMS-7	GMS-9	GMS-11
GMS-1	1.20%	58.40%	60.50%	58%	61%	44.40%
GMS-3		5.50%	65.80%	67.10%	68.80%	63.90%
GMS-5			4.00%	70.10%	69.70%	60.60%
GMS-7				3.20%	68.50%	66.60%
GMS-9					2.80%	61.20%
GMS-11						15.20%
Total genes detected	342	531	504	539	494	746
Simpson's (1/D)	121.9	185.4	199	168.2	179.7	204.9
Shannon Weaver's H'	5.216	5.294	5.365	5.355	5.336	5.392
Shannon Weaver's evenness	0.801	0.813	0.824	0.823	0.82	0.829

- More genes were detected in the deeper sediments;
- The communities in deep sediments were more diverse;
- Less overlap was found between samples within a station as distance increased.

Table 5. Percentage of Genes Overlapping between Similar Depth Increments between the Stations as Well as Overall

	GMS (overall)	GMS-1	GMS-3	GMS-5	GMS-7	GMS-9	GMS-11
GMT (overall)	68.5						
GMT-1		72.0	52.5	51.7	49.6	53.3	37.7
GMT-3		68.6	61.1	62.7	59.7	63.9	47.8
GMT-5		68.2	56.1	57.7	54.1	59.1	42.1
GMT-7		66.9	63.7	61.5	63.7	63.5	51.6
GMT-9		56.8	65.3	69.3	71.7	68.2	62.8
GMT-11		51.3	66.3	62.3	65.5	64.9	64.09

RESULTS: Cluster Analysis.

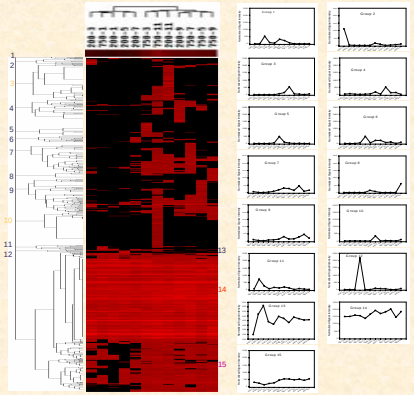

Figure 1. Clusters of all genes. 15 patterns were observed. The most obvious pattern was Group 14 (223/23.9 %) whose members were abundant across all of the samples. The next obvious group is Group 15 (9.8%), the members of which were most from the deeper layers. Group 3 and 10 existed in low abundance and to be unique to the deepest samples. Other groups were present in low abundance across most samples and may be highly abundant in one to a few of the sediment samples.

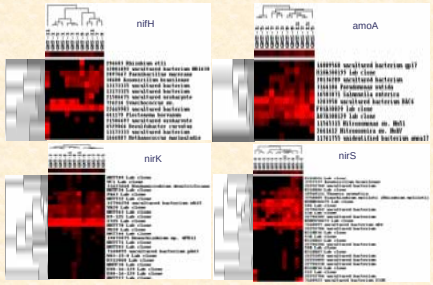

Figure 2. Clusters of nitrogen cycling genes. The majority of the *amo*A genes observed were environmental clones. For instance, the three genes, 3283950, A07A300129 and F01A30029, from environments, were observed across all of samples, and were most abundant. Most of the *nif*H, *nir*S and *nir*K genes observed were also from environmental clones rather than from pure culture strains.

RESULTS: DCA analysis.

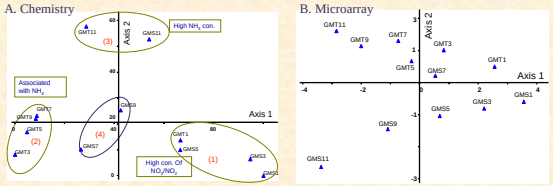

Figure 3. DCA analysis. Four distinct clusters can be visualized from the DCA analysis of geochemical parameters (A). The DCA analysis of the hybridization data demonstrates an interesting divergence that does not closely reflect the environmental DCA. The shallow samples of both stations are closely correlated. As the sediment depth increase, the samples separate. The near-neighbor samples were closely related; and as the depth increased, the samples became less similar to the shallow layers as well as being less similar between stations (B).

Table 6. Mantel Test for Station 3

		Number of genes significantly correlated to geochemistry					
Probe Categories	Total	NH_4	NO_3	SO_4reduction	Depth	Porosity	Si
Carbon degradation							
Chitinase	27	4 (0.26 - 0.38)	2 (0.87 - 0.98)	nd	5 (0.58 - 0.84)	7 (0.58 - 0.80)	8 (0.22 - 0.39)
Endoglucanase	38	5 (0.54 - 0.98)	4 (0.62 - 0.77)	nd	8 (0.56 - 0.92)	10 (0.23 - 0.91)	8 (0.19 - 0.52)
Laccase	16	0	3 (0.72 - 0.84)	nd	0	0	0
mdh	2	0	1 (0.96)	nd	0	0	0
mmo	10	2 (0.28 - 0.84)	0	nd	0	0	0
mmr	2	0	0	nd	0	0	0
Xylanase	39	4 (0.45 - 0. 86)	1 (0.52)	nd	6 (0.37 - 0.92)	6 (0.29 - 0.62)	6 (0.43 - 0.64)
Carbon fixation							
cbbL	4	0	1 (0.69)	nd	0	0	0
cbbM	7	0	0	nd	0	0	0
FTHFS	11	1 (0.72)	0	nd	0	0	0
rbcL	2	0	0	nd	0	0	0
Contaminant degradation	418	39 (0.25 - 0.95)	28 (0.27 - 0.97)	nd	85 (0.31 - 0.93)	93 (0.14 - 0.96)	61 (0.31 - 0.85)
Metal tolerance	28	6 (0.28 - 0.78)	3 (0.31 - 0.83)	nd	0	0	0
Nitrogen cycle							
Urease	2	0	0	nd	nd	nd	nd
nasA	6	1 (0.62)	0	nd	nd	nd	nd
nosZ	5	0	1 (0.34)	nd	nd	nd	nd
napA	1	0	0	nd	nd	nd	nd
narB	1	0	1 (0.28)	nd	nd	nd	nd
narG	10	2 (0.60 - 0.95)	0	nd	nd	nd	nd
nor	6	0	0	nd	nd	nd	nd
amoA	9	5 (0.34 - 0.86)	1 (0.30)	nd	nd	nd	nd
amoK	19	2 (0.88 - 0.95)	0	nd	nd	nd	nd
nirS	19	1 (0.69)	0	nd	nd	nd	nd
nifD	7	1 (0.52)	0	nd	nd	nd	nd
nifH	12	2 (0.42 - 0.65)	1 (0.29)	nd	nd	nd	nd
Phosphorus metabolism	19	3 (0.47 - 0.90)	0	nd	0	0	0
Sulfur reduction							
apr-sulfur	8	1 (0.42)	0	0	1 (0.27)	1 (0.40)	nd
dsr-sulfur	40	4 (0.40 - 0.88)	6 (0.27 - 0.99)	7 (0.28 - 0.99)	8 (0.42 - 0.82)	10 (0.46 - 0.86)	nd

Although no overall correlation among the communities and environmental variables was found; mantel tests revealed that many of the individual functional group of genes were correlated to the measured environmental variables.

CONCLUTIONS

- Microbial communities differ both vertically and horizontally. There exists biogeographical distributions that are nonrandom associations.
- The correlation between the microbial communities present and the environmental variables was weak.
- Diversity and richness of functional genes increased with depth.
- Individual functional groups of genes were correlated to the measured environmental variables.

ACKNOWLEDGEMENT

This research was funded by the U.S. Department of Energy (Office of Biological and Environmental Research, Office of Science) grants from the Genomes To Life Program and the ERSP Program.